Chapter 1

Introduction

One of our goals in this book is to equip the reader with a unifying view of linear algebra, or at least of what is traditionally studied under this name in university courses. With this mission in mind, we start with a *preview* of the subject, and describe its main achievements in lay terms.

To begin with a few words of praise: linear algebra is a very simple and useful subject, underlying most of other areas of mathematics, as well as its applications to physics, computer science, engineering, and economics. What makes linear algebra useful and efficient is that it provides ultimate solutions to several important mathematical problems. Furthermore, as should be expected of a truly fruitful mathematical theory, the problems it solves can be formulated in a rather elementary language, and make sense even before any advanced machinery is developed. Even better, the *answers* to these problems can also be described in elementary terms (in contrast with the *justification* of those answers, which better be postponed until adequate tools are developed). Finally, those several problems we are talking about are similar in their nature; namely, they all have the form of problems of *classification* of very basic mathematical objects.

Yet unready to discuss the general idea of classification in mathematics, we start off with a geometric introduction to vectors, and a summary of compex numbers. Then we work out a non-trivial model example: classification of quadratic curves on the plane. Then, with this example in mind, we will be able to say a few general words about the idea of classification in general, and then present in elementary, down-to-earth terms the main problems of linear algebra, and the answers to these problems. At that point, the layout of the further material will also become clear.

1 Vectors in Geometry

Operations and their properties

The following definition of vectors can be found in elementary geometry textbooks, see for instance [4].

A **directed segment** $\overrightarrow{AB}$ on the plane or in space is specified by an ordered pair of points: the **tail** A and the **head** B. Two directed segments $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are said to **represent the same vector** if they are obtained from one another by translation. In other words, the lines AB and CD must be parallel, the lengths $|AB|$ and $|CD|$ must be equal, and the segments must point toward the same of the two possible directions (Figure 1).

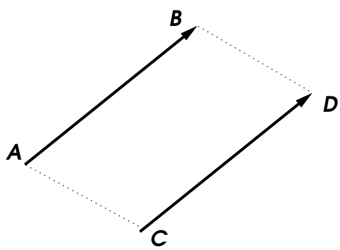

Figure 1

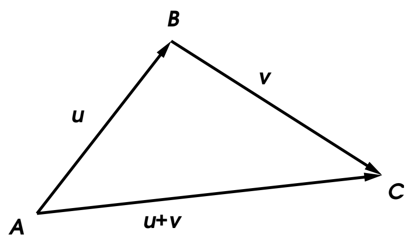

Figure 2

A trip from A to B followed by a trip from B to C results in a trip from A to C. This observation motivates the definition of the **vector sum** $\mathbf{w} = \mathbf{v} + \mathbf{u}$ of two vectors $\mathbf{v}$ and $\mathbf{u}$: if $\overrightarrow{AB}$ represents $\mathbf{v}$ and $\overrightarrow{BC}$ represents $\mathbf{u}$ then $\overrightarrow{AC}$ represents their sum $\mathbf{w}$ (Figure 2).

The vector $3\mathbf{v} = \mathbf{v} + \mathbf{v} + \mathbf{v}$ has the same direction as $\mathbf{v}$ but is 3 times longer. Generalizing this example one arrives at the definition of the **multiplication of a vector by a scalar**: given a vector $\mathbf{v}$ and a real number α, the result of their multiplication is a vector, denoted $\alpha\mathbf{v}$, which has the same direction as $\mathbf{v}$ but is α times longer. The last phrase calls for comments since it is literally true only for $\alpha > 1$. If $0 < \alpha < 1$, being "α times longer" actually means "shorter." If $\alpha < 0$, the direction of $\alpha\mathbf{v}$ is in fact opposite to the direction of $\mathbf{v}$. Finally, $0\mathbf{v} = \mathbf{0}$ is the **zero vector** represented by directed segments $\overrightarrow{AA}$ of zero length.

Combining the operations of vector addition and multiplication by scalars we can form expressions $\alpha\mathbf{u}+\beta\mathbf{v}+...+\gamma\mathbf{w}$. They are called

linear combinations of the vectors $\mathbf{u}, \mathbf{v}, ..., \mathbf{w}$ with the coefficients $\alpha, \beta, ..., \gamma$.

The pictures of a parallelogram and parallelepiped (Figures 3 and 4) prove that the addition of vectors is **commutative** and **associative**: for all vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$,

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} \quad \text{and} \quad (\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}).$$

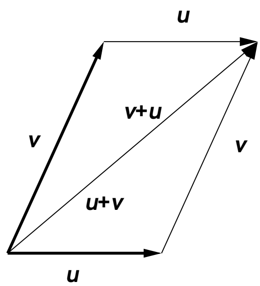

Figure 3

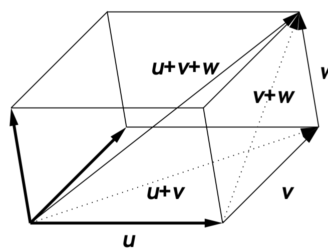

Figure 4

From properties of proportional segments and similar triangles, the reader will easily derive the following two **distributive laws**: for all vectors $\mathbf{u}, \mathbf{v}$ and scalars α, β,

$$(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u} \quad \text{and} \quad \alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}.$$

EXERCISES

1. A mass m rests on an inclined plane making 30° with the horizontal plane. Find the forces of friction and reaction acting on the mass. ✓

2. A ferry, capable of making 5 *mph*, shuttles across a river of width 0.6 *mi* with a strong current of 3 *mph*. How long does each round trip take? ✓

3. Prove that for every closed broken line $ABC \dots DE$,

$$\overrightarrow{AB} + \overrightarrow{BC} + \cdots + \overrightarrow{DE} + \overrightarrow{EA} = \mathbf{0}.$$

4. Three medians of a triangle ABC intersect at one point M called the **barycenter** of the triangle. Let O be any point on the plane. Prove that

$$\overrightarrow{OM} = \frac{1}{3}(\overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC}).$$

5. Prove that $\overrightarrow{MA} + \overrightarrow{MB} + \overrightarrow{MC} = \mathbf{0}$ if and only if M is the barycenter of the triangle ABC.

6.* Along three circles lying in the same plane, the vertices of a triangle are moving clockwise with the equal constant angular velocities. Find how the barycenter of the triangle is moving. ✓

7. Prove that if AA' is a median in a triangle ABC, then

$$\overrightarrow{AA'} = \frac{1}{2}(\overrightarrow{AB} + \overrightarrow{AC}).$$

8. Prove that from medians of a triangle, another triangle can be formed. ↯

9. Sides of one triangle are parallel to the medians of another. Prove that the medians of the latter triangle are parallel to the sides of the former one.

10. From medians of a given triangle, a new triangle is formed, and from its medians, yet another triangle is formed. Prove that the third triangle is similar to the first one, and find the coefficient of similarity. ✓

11. Midpoints of AB and CD, and of BC and DE are connected by two segments, whose midpoints are also connected. Prove that the resulting segment is parallel to AE and congruent to $AE/4$.

12. Prove that a point X lies on the segment AB if and only if for any origin O and some scalar $0 \leq \lambda \leq 1$ the radius-vector $\overrightarrow{OX}$ has the form:

$$\overrightarrow{OX} = \lambda\overrightarrow{OA} + (1-\lambda)\overrightarrow{OB}.$$

13.* Given a triangle ABC, we construct a new triangle $A'B'C'$ in such a way that A' is centrally symmetric to A with respect to the center B, B' centrally symmetric to B with respect to C, and C' centrally symmetric to C with respect to A, and then erase the original triangle. Reconstruct ABC from $A'B'C'$ by straightedge and compass. ✓

Coordinates

From a point O in space, draw three directed segments $\overrightarrow{OA}$, $\overrightarrow{OB}$, and $\overrightarrow{OC}$ not lying in the same plane and denote by $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ the vectors they represent. Then every vector $\mathbf{u} = \overrightarrow{OU}$ can be uniquely written as a linear combination of $\mathbf{i}, \mathbf{j}, \mathbf{k}$ (Figure 5):

$$\mathbf{u} = \alpha\mathbf{i} + \beta\mathbf{j} + \gamma\mathbf{k}.$$

The coefficients form the array (α, β, γ) of **coordinates** of the vector $\mathbf{u}$ (and of the point U) with respect to the **basis** $\mathbf{i}, \mathbf{j}, \mathbf{k}$ (or the **coordinate system** $OABC$).

Multiplying $\mathbf{u}$ by a scalar λ or adding another vector $\mathbf{u}' = \alpha'\mathbf{i} + \beta'\mathbf{j} + \gamma'\mathbf{k}$, and using the above algebraic properties of the operations with vectors, we find:

$$\lambda\mathbf{u} = \lambda\alpha\mathbf{i} + \lambda\beta\mathbf{j} + \lambda\gamma\mathbf{k}, \text{ and } \mathbf{u} + \mathbf{u}' = (\alpha+\alpha')\mathbf{i} + (\beta+\beta')\mathbf{j} + (\gamma+\gamma')\mathbf{k}.$$

Thus, the geometric operations with vectors are expressed by componentwise operations with the arrays of their coordinates:

$$\lambda(\alpha,\beta,\gamma) = (\lambda\alpha,\lambda\beta,\lambda\gamma),$$
$$(\alpha,\beta,\gamma)+(\alpha',\beta',\gamma') = (\alpha+\alpha',\beta+\beta',\gamma+\gamma').$$

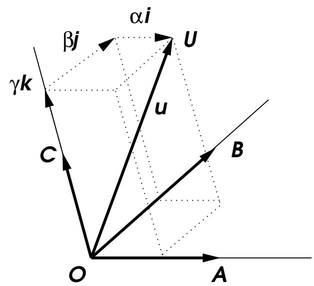

Figure 5

What is a vector?

No doubt, the idea of vectors is not new to the reader. However, some subtleties of the above introduction do not easily meet the eye, and we would like to say here a few words about them.

As many other mathematical notions, vectors come from physics, where they represent quantities, such as velocities and forces, which are characterized by their magnitude and direction. Yet, the popular slogan "Vectors *are* magnitude and direction" does not qualify for a mathematical definition of vectors, e.g. because it does not tell us how to operate with them.

The computer science definition of vectors as *arrays* of numbers, to be added "apples with apples, oranges with oranges," will meet the following objection by physicists. When a coordinate system rotates, the coordinates of the *same* force or velocity will change, but the numbers of apples and oranges won't. Thus forces and velocities *are not* arrays of numbers.

The geometric notion of a directed segment resolves this problem. Note however, that calling directed segments *vectors* would constitute abuse of terminology. Indeed, strictly speaking, directed segments can be added only when the head of one of them coincides with the tail of the other.

So, what is a vector? In our formulations, we actually avoided answering this question directly, and said instead that *two directed segments represent the same vector if ...* Such wording is due to pedagogical wisdom of the authors of elementary geometry textbooks, because a direct answer sounds quite abstract: *A vector is the class of all directed segments obtained from each other by translation in space.* Such a class is shown in Figure 6.

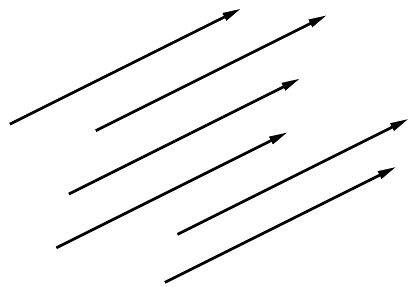

Figure 6

This picture has another interpretation: For every point in space (the tail of an arrow), it indicates a new position (the head). The geometric transformation in space defined this way is translation. This leads to another attractive point of view: a vector *is* a translation. Then the sum of two vectors is the *composition* of the translations.

The dot product

This operation encodes *metric* concepts of elementary Euclidean geometry, such as lengths and angles. Given two vectors $\mathbf{u}$ and $\mathbf{v}$ of lengths $|\mathbf{u}|$ and $|\mathbf{v}|$ and making the angle θ to each other, their dot product (also called **inner product** or **scalar product**) is a *number* defined by the formula:

$$\langle \mathbf{u}, \mathbf{v} \rangle = |\mathbf{u}|\ |\mathbf{v}|\ \cos\theta.$$

Of the following properties, the first three are easy (check them!):

(a) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ (**symmetricity**);

(b) $\langle \mathbf{u}, \mathbf{u} \rangle = |\mathbf{u}|^2 > 0$ unless $\mathbf{u} = \mathbf{0}$ (**positivity**);

(c) $\langle \lambda\mathbf{u}, \mathbf{v} \rangle = \lambda\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, \lambda\mathbf{v} \rangle$ (**homogeneity**);

(d) $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ (**additivity** with respect to the first factor).

To prove the last property, note that due to homogeneity, it suffices to check it assuming that $\mathbf{w}$ is a **unit vector**, i.e. $|\mathbf{w}| = 1$. In this case, consider (Figure 7) a triangle ABC such that $\overrightarrow{AB} = \mathbf{u}$, $\overrightarrow{BC} = \mathbf{v}$, and therefore $\overrightarrow{AC} = \mathbf{u} + \mathbf{v}$, and let $\overrightarrow{OW} = \mathbf{w}$. We can consider the line OW as the number line, with the points O and W representing the numbers 0 and 1 respectively, and denote by α, β, γ the *numbers* representing perpendicular projections to this line of the vertices A, B, C of the triangle. Then

$$\langle \overrightarrow{AB}, \mathbf{w} \rangle = \beta - \alpha, \ \langle \overrightarrow{BC}, \mathbf{w} \rangle = \gamma - \beta, \ \text{and} \langle \overrightarrow{AC}, \mathbf{w} \rangle = \gamma - \alpha.$$

The required identity follows, because $\gamma - \alpha = (\gamma - \beta) + (\beta - \alpha)$.

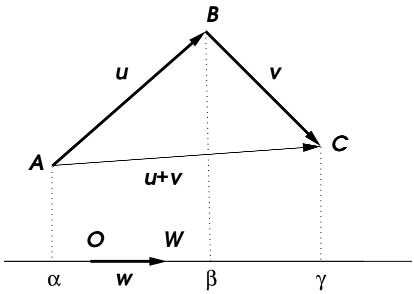

Figure 7

Combining the properties (c) and (d) with (a), we obtain the following identities, expressing **bilinearity** of the dot product (i.e. linearity with respect to each factor):

$$\langle \alpha \mathbf{u} + \beta \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle$$

$$\langle \mathbf{w}, \alpha \mathbf{u} + \beta \mathbf{v} \rangle = \alpha \langle \mathbf{w}, \mathbf{u} \rangle + \beta \langle \mathbf{w}, \mathbf{v} \rangle.$$

The following example illustrates the use of nice algebraic properties of dot product in elementary geometry.

`Example.` Given a triangle ABC, let us denote by $\mathbf{u}$ and $\mathbf{v}$ the vectors represented by the directed segments $\overrightarrow{AB}$ and $\overrightarrow{AC}$ and use properties of the inner product in order to compute the length $|BC|$. Notice that the segment $\overrightarrow{BC}$ represents $\mathbf{v} - \mathbf{u}$. We have:

$$\begin{aligned} |BC|^2 = \langle \mathbf{v} - \mathbf{u}, \mathbf{v} - \mathbf{u} \rangle &= \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{u} \rangle - 2 \langle \mathbf{u}, \mathbf{v} \rangle \\ &= |AC|^2 + |AB|^2 - 2|AB|\ |AC|\ \mathbf{cos}\, \theta. \end{aligned}$$

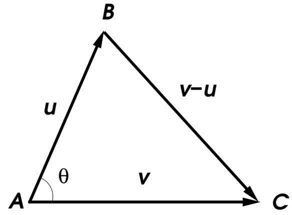

Figure 8

This is the famous **Law of Cosines** in trigonometry.

When the vectors **u** and **v** are **orthogonal**, i.e. $\langle \mathbf{u}, \mathbf{v}\rangle = 0$, then the formula turns into the **Pythagorean theorem**:

$$|\mathbf{u} \pm \mathbf{v}|^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2.$$

When basis vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are pairwise orthogonal and *unit*, the coordinate system is called **Cartesian**.[1] We have:

$$\langle \mathbf{i}, \mathbf{i}\rangle = \langle \mathbf{j}, \mathbf{j}\rangle = \langle \mathbf{k}, \mathbf{k}\rangle = 1, \text{ and } \langle \mathbf{i}, \mathbf{j}\rangle = \langle \mathbf{j}, \mathbf{k}\rangle = \langle \mathbf{k}, \mathbf{i}\rangle = 0.$$

Thus, in Cartesian coordinates, the inner squares and the dot products of vectors $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ and $\mathbf{r}' = x'\mathbf{i} + y'\mathbf{j} + z'\mathbf{k}$ are given by the formulas:

$$|\mathbf{r}|^2 = x^2 + y^2 + z^2, \quad \langle \mathbf{r}, \mathbf{r}'\rangle = xx' + yy' + zz'.$$

EXERCISES

14. Prove the **Cauchy – Schwarz inequality**: $\langle \mathbf{u}, \mathbf{v}\rangle^2 \le \langle \mathbf{u}, \mathbf{u}\rangle\langle \mathbf{v}, \mathbf{v}\rangle$. In which cases does the inequality turn into equality? Deduce the **triangle inequality**: $|\mathbf{u} + \mathbf{v}| \le |\mathbf{u}| + |\mathbf{v}|$. ↯

15. Compute the inner product $\langle \overrightarrow{AB}, \overrightarrow{BC}\rangle$ if ABC is a regular triangle inscribed into a unit circle. ✓

16. Prove that if the sum of three unit vectors is equal to **0**, then the angle between each pair of these vectors is equal to $120°$.

17. Express the inner product $\langle \mathbf{u}, \mathbf{v}\rangle$ in terms of the lengths $|\mathbf{u}|, |\mathbf{v}|, |\mathbf{u} + \mathbf{v}|$ of the two vectors and of their sum. ✓

18. (a) Prove that if four unit vectors lying in the same plane add up to **0**, then they form two pairs of opposite vectors. (b) Does this remain true if the vectors do not have to lie in the same plane? ✓

[1] After René **Descartes** (1596–1650).

19.* Let $AB \ldots E$ be a regular polygon with the center O. Prove that

$$\overrightarrow{OA} + \overrightarrow{OB} + \cdots + \overrightarrow{OE} = \mathbf{0}. \qquad ↯$$

20. Prove that if $\mathbf{u} + \mathbf{v}$ and $\mathbf{u} - \mathbf{v}$ are perpendicular, then $|\mathbf{u}| = |\mathbf{v}|$.

21. For arbitrary vectors $\mathbf{u}$ and $\mathbf{v}$, verify the equality:

$$|\mathbf{u} + \mathbf{v}|^2 + |\mathbf{u} - \mathbf{v}|^2 = 2|\mathbf{u}|^2 + 2|\mathbf{v}|^2,$$

and derive the theorem: The sum of the squares of the diagonals of a parallelogram is equal to the sum of the squares of the sides.

22. Prove that for every triangle ABC and every point X in space,

$$\overrightarrow{XA} \cdot \overrightarrow{BC} + \overrightarrow{XB} \cdot \overrightarrow{CA} + \overrightarrow{XC} \cdot \overrightarrow{AB} = 0. \qquad ↯$$

23.* For four arbitrary points A, B, C, and D in space, prove that if the lines AC and BD are perpendicular, then $AB^2 + CD^2 = BC^2 + DA^2$, and *vice versa.* ↯

24.* Given a quadrilateral with perpendicular diagonals. Show that every quadrilateral, whose sides are respectively congruent to the sides of the given one, has perpendicular diagonals. ↯

25.* A regular triangle ABC is inscribed into a circle of radius R. Prove that for every point X of this circle, $XA^2 + XB^2 + XC^2 = 6R^2$. ↯

26.* Let $A_1B_1A_2B_2 \ldots A_nB_n$ be a $2n$-gon inscribed into a circle. Prove that the length of the vector $\overrightarrow{A_1B_1} + \overrightarrow{A_2B_2} + \cdots + \overrightarrow{A_nB_n}$ does not exceed the diameter. ↯

27.* A polyhedron is filled with air under pressure. The pressure force to each face is the vector perpendicular to the face, proportional to the area of the face, and directed to the exterior of the polyhedron. Prove that the sum of these vectors is equal to $\mathbf{0}$. ↯

2 Complex Numbers

Law and Order

Life is unfair: The quadratic equation $x^2 - 1 = 0$ has two solutions $x = \pm 1$, but a similar equation $x^2 + 1 = 0$ has no solutions at all. To restore justice one introduces new number i, the **imaginary unit**, such that $i^2 = -1$, and thus $x = \pm i$ become two solutions to the equation. This is how complex numbers could have been invented.

More formally, complex numbers are introduced as ordered pairs (a, b) of real numbers, written in the form $z = a + bi$. The real numbers a and b are called respectively the **real part** and **imaginary part** of the complex number z, and are denoted $a = \operatorname{Re} z$ and $b = \operatorname{Im} z$.

The sum of $z = a + bi$ and $w = c + di$ is defined as

$$z + w = (a + c) + (b + d)i.$$

The product is defined so as to comply with the relation $i^2 = -1$:

$$zw = ac + bdi^2 + adi + bci = (ac - bd) + (ad + bc)i.$$

The operations of addition and multiplication of complex numbers enjoy the same properties as those of real numbers do. In particular, the product is commutative and associative.

The complex number $\bar{z} = a - bi$ is called **complex conjugate** to $z = a + bi$. The operation of complex conjugation *respects* sums and products:

$$\overline{z + w} = \bar{z} + \bar{w} \quad \text{and} \quad \overline{zw} = \bar{z}\bar{w}.$$

This can be easily checked from definitions, but there is a more profound explanation. The equation $x^2 + 1 = 0$ has two roots, i and $-i$, and the choice of the one to be called i is totally ambiguous. The complex conjugation consists in systematic renaming i by $-i$ and *vice versa*, and such renaming cannot affect properties of complex numbers.

Complex numbers satisfying $\bar{z} = z$ are exactly the real numbers $a + 0i$. We will see that this point of view on real numbers as complex numbers *invariant* under complex conjugation is quite fruitful.

The product $z\bar{z} = a^2 + b^2$ (check this formula!) is real, and is positive unless $z = 0 + 0i = 0$. This shows that

$$\frac{1}{z} = \frac{\bar{z}}{z\bar{z}} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} i.$$

Hence the division by z is well-defined for any non-zero complex number z. In terminology of Abstract Algebra, complex numbers form therefore a **field**[2] (just as real or rational numbers do).

The field of complex numbers is denoted by $\mathbb{C}$ (while $\mathbb{R}$ stands for reals, and $\mathbb{Q}$ for rationals).

The non-negative real number $|z| = \sqrt{z\bar{z}} = \sqrt{a^2+b^2}$ is called the **absolute value** of z. The absolute value function is **multiplicative**:

$$|zw| = \sqrt{zw\overline{zw}} = \sqrt{z\bar{z}w\bar{w}} = |z| \cdot |w|.$$

It actually coincides with the absolute value of real numbers when applied to complex numbers with zero imaginary part: $|a+0i| = |a|$.

EXERCISES

28. Can complex numbers be: real? real *and* imaginary? neither? ✓

29. Compute: (a) $(1+i)/(3-2i)$; (b) $(\mathbf{cos}\,\pi/3 + i\,\mathbf{sin}\,\pi/3)^{-1}$. ✓

30. Verify the commutative and distributive laws for multiplication of complex numbers.

31. Show that z^{-1} is real proportional to $\bar{z}$ and find the proportionality coefficient. ✓

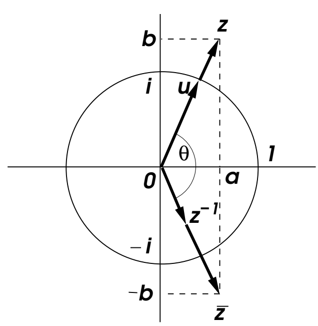

Figure 9

Geometry

We can identify complex numbers $z = a + bi$ with points (a, b) on the real coordinate plane (Figure 9). This way, the number 0 is identified with the origin, and 1 and i become the unit basis vectors

[2]This requires that a set be equipped with commutative and associative operations (called addition and multiplication) satisfying the **distributive law** $z(v+w) = zv + zw$, possessing the zero and unit elements 0 and 1, additive opposites $-z$ for every z, and multiplicative inverses $1/z$ *for every* $z \neq 0$.

(1, 0) and (0, 1). The coordinate axes are called respectively the real and imaginary axes. Addition of complex numbers coincides with the operation of vector sum (Figure 10).

The absolute value function has the geometrical meaning of the distance from the origin: $|z| = \langle z, z\rangle^{1/2}$. In particular, the triangle inequality $|z + w| \leq |z| + |w|$ holds true. Complex numbers of unit absolute value $|z| = 1$ form the unit circle centered at the origin.

The operation of complex conjugation acts on the radius-vectors z as the reflection about the real axis.

In order to describe a geometric meaning of complex multiplication, let us study the way multiplication by a given complex number z acts on all complex numbers w, i.e. consider the function $w \mapsto zw$. For this, write the vector representing a non-zero complex number z in the **polar** (or trigonometric) form $z = ru$ where $r = |z|$ is a positive real number, and $u = z/|z| = \mathbf{cos}\,\theta + i\,\mathbf{sin}\,\theta$ has absolute value 1. Here $\theta = \mathbf{arg}\, z$, called the **argument** of the complex number z, is the angle that z as a vector makes with the positive direction of the real axis.

Clearly, multiplication by r acts on all vectors w by stretching them r times. Multiplication by u applied to $w = x + yi$ yields a new complex number $uw = X + Yi$ according to the rule:

$$X = \text{Re}\ [(\mathbf{cos}\,\theta + i\,\mathbf{sin}\,\theta)(x + yi)] = x\,\mathbf{cos}\,\theta - y\,\mathbf{sin}\,\theta$$
$$Y = \text{Im}\ [(\mathbf{cos}\,\theta + i\,\mathbf{sin}\,\theta)(x + yi)] = x\,\mathbf{sin}\,\theta + y\,\mathbf{cos}\,\theta.$$

We claim that this is the result of counter-clockwise rotation of w through the angle θ.

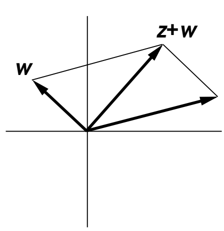

Figure 10

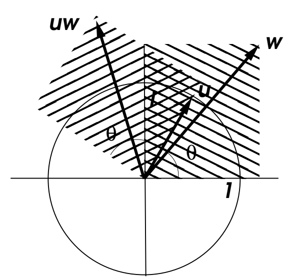

Figure 11

Indeed, $w = x + yi$ is represented by a diagonal of the rectangle with the sides $x \times 1$ and $y \times i$. The rotated rectangle is still a rectangle, whose sides are $x(\mathbf{cos}\,\theta + i\,\mathbf{sin}\,\theta)$ and $y(-\,\mathbf{sin}\,\theta + i\,\mathbf{cos}\,\theta)$

(since $-\sin\theta + i\cos\theta$ is the result of the counter-clockwise rotation of i thrrough the angle θ). Therefore diagonal of the latter rectangle coincides with $X + Yi = uw$.

In Section 3 we will encounter a similar computation, with the following notable difference. Here we rotate vectors, while the coordinate system remains unchanged. In Section 3, we will rotate the coordinate system from the old to a new position, and express old coordinates of a vector via new coordinates of the *same* vector.

Anyway, the conclusion is that multiplication by z is the composition of two operations: stretching $|z|$ times, and rotating through the angle $\arg z$.

In other words, the product operation of two complex numbers sums their arguments and multiplies absolute values:

$$|zw| = |z| \cdot |w|, \quad \arg zw = \arg z + \arg w \text{ modulo } 2\pi.$$

For example, if $z = r(\cos\theta + i\sin\theta)$, then $z^n = r^n(\cos n\theta + i\sin n\theta)$.

EXERCISES

32. Sketch the solution set to the following system of inequalities:

$$|z-1| \leq 1, \quad |z| \leq 1, \quad \mathrm{Re}(iz) \leq 0.$$

33. Compute absolute values and arguments of (a) $1-i$, (b) $1-i\sqrt{3}$. ✓

34. Compute $\left(\frac{\sqrt{3}+i}{2}\right)^{100}$. ✓

35. Express $\cos 3\theta$ and $\sin 3\theta$ in terms of $\cos\theta$ and $\sin\theta$. ϟ

36. Express $\cos(\theta_1 + \theta_2)$ and $\sin(\theta_1 + \theta_2)$ in terms of $\cos\theta_i$ and $\sin\theta_i$.

The Fundamental Theorem of Algebra

A degree 2 polynomial $z^2 + pz + q$ has two roots

$$z_\pm = \frac{-p \pm \sqrt{p^2 - 4q}}{2}.$$

This **quadratic formula** works regardless of the sign of the **discriminant** $p^2 - 4q$, provided that we allow the roots to be complex, and take in account **multiplicity**. Namely, if $p^2 - 4q = 0$, $z^2 + pz + q = (z + p/2)^2$ and therefore the single root $z = -p/2$ has multiplicity two. If $p^2 - 4q < 0$ the roots are complex conjugate with $\mathrm{Re}\, z_\pm = -p/2$, $\mathrm{Im}\, z_\pm = \pm\sqrt{|p^2 - 4q|}/2$. The Fundamental Theorem of Algebra shows that not only justice has been restored, but that any degree n polynomial has n complex roots, possibly — multiple.

`Theorem.` ***A degree*** n ***polynomial***

$$P(z) = z^n + a_1 z^{n-1} + ... + a_{n-1} z + a_n$$

with complex coefficients $a_1, ..., a_n$ ***factors as***

$$P(z) = (z - z_1)^{m_1} ... (z - z_r)^{m_r}.$$

Here $z_1, ..., z_r$ ***are complex roots of*** P***, and*** $m_1, ..., m_r$ ***their multiplicities,*** $m_1 + \cdots + m_r = n$***.***

A proof of this theorem deserves a separate chapter (if not a book). Many proofs are known, based on various ideas of Algebra, Analysis or Topology. We refer to [6] for an exposition of the classical proof due to Euler, Lagrange and de Foncenex, which is almost entirely algebraic. Here we merely illustrate the theorem with several examples.

`Examples.` (a) To solve the quadratic equation $z^2 = w$, equate the absolute value r and argument θ of the given complex number w with those of z^2:

$$|z|^2 = \rho, \quad 2\,\mathbf{arg}\, z = \phi + 2\pi k, \quad k = 0, \pm 1, \pm 2, \ldots.$$

We find: $|z| = \sqrt{\rho}$, and $\mathbf{arg}\, z = \phi/2 + \pi k$. Increasing $\mathbf{arg}\, z$ by even multiples π does not change z, and by odd changes z to $-z$. Thus the equation has two solutions:

$$z = \pm\sqrt{\rho}\left(\mathbf{cos}\,\frac{\phi}{2} + i\,\mathbf{sin}\,\frac{\phi}{2}\right).$$

(b) The equation $z^2 + pz + q = 0$ with coefficients $p, q \in \mathbb{C}$ has two complex solutions given by the quadratic formula (see above), because according to Example (a), the **square root** of a complex number takes on two opposite values (distinct, unless both are equal to 0).

(c) The complex numbers $1, i, -1, -i$ are the roots of the polynomial $z^4 - 1 = (z^2 - 1)(z^2 + 1) = (z - 1)(z + 1)(z - i)(z + i)$.

(d) There are n complex n**th roots of unity** (see Figure 12, where $n = 5$). Namely, if $z = r(\mathbf{cos}\,\theta + i\,\mathbf{sin}\,\theta)$ satisfies $z^n = 1$, then $r^n = 1$ (and hence $r = 1$), and $n\theta = 2\pi k$, $k = 0, \pm 1, \pm 2, ...$ Therefore $\theta = 2\pi k/n$, where only the remainder of k modulo n is relevant. Thus the n roots are:

$$z = \mathbf{cos}\,\frac{2\pi k}{n} + i\,\mathbf{sin}\,\frac{2\pi k}{n}, \;\; k = 0, 1, 2, ..., n - 1.$$

For instance, if $n = 3$, the roots are 1 and

$$\cos\frac{2\pi}{3} \pm i\sin\frac{2\pi}{3} = -\frac{1}{2} \pm i\frac{\sqrt{3}}{2}.$$

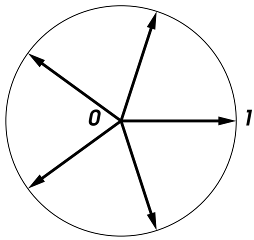

Figure 12

As illustrated by the previous two examples, even if all coefficients $a_1, \ldots, a_n$ of a polynomial P are real, its roots don't have to be real. But then the non-real roots come in pairs of complex conjugate ones. To verify this, we can use the fact that being real means stay *invariant* (i.e. unchanged) under complex conjugation. Namely, $\bar{a}_i = a_i$ for all i means that

$$\overline{P(\bar{z})} = z^n + \bar{a}_1 z^{n-1} + ... + \bar{a}_n = P(z).$$

Therefore we have two factorizations of the same polynomial:

$$\bar{P}(\bar{z}) = (z - \bar{z}_1)^{m_1} ... (z - \bar{z}_r)^{m_r} = (z - z_1)^{m_1} ... (z - z_r)^{m_r} = P(z).$$

They can differ only by orders of the factors. Thus, for each non-real root z_i of P, the complex conjugate $\bar{z}_i$ must be also a root, and of the same multiplicity.

Expanding the product

$$(z - z_1)...(z - z_n) = z^n - (z_1 + ... + z_n)z^{n-1} + ... + (-1)^n z_1 ... z_n$$

we can express coefficients $a_1, ..., a_n$ of the polynomial in terms of the roots $z_1, ..., z_n$ (here multiple roots are repeated according to their multiplicities). In particular, the sum and product of the roots are

$$z_1 + ... + z_n = -a_1, \quad z_1 ... z_n = (-1)^n a_n.$$

These relations generalize **Vieta's theorem** $z_+ + z_- = -p$, $z_+ z_- = q$ about roots $z_\pm$ of quadratic equations $z^2 + pz + q = 0$.

EXERCISES

37. Prove **Bézout's theorem**[3]: A number z_0 is a root of a polynomial P in one variable z if and only if P is divisible by $z - z_0$. ↯

38. Find roots of degree 2 polynomials:

$$z^2 - 4z + 5, \quad z^2 - iz + 1, \quad z^2 - 2(1+i)z + 2i, \quad z^2 - 2z + i\sqrt{3}. \qquad \checkmark$$

39. Find all roots of polynomials:

$$z^3 + 8, \quad z^3 + i, \quad z^4 + 4z^2 + 4, \quad z^4 - 2z^2 + 4, \quad z^6 + 1. \qquad \checkmark$$

40. Prove that every polynomial with real coefficients factors into the product of polynomials of degree 1 and 2 with real coefficients. ↯

41. Prove that the sum of all 5th roots of unity is equal to 0. ↯

42.* Find general **Vieta's formulas** [4] expressing all coefficients of a polynomial in terms of its roots. ✓

The Exponential Function

Consider the series

$$1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \ldots + \frac{z^n}{n!} + \ldots$$

Applying the ratio test for convergence of infinite series,

$$\left| \frac{z^n (n-1)!}{n! z^{n-1}} \right| = \frac{|z|}{n} \to 0 < 1 \text{ as } n \to \infty,$$

we conclude that the series converges absolutely for any complex number z. The sum of the series is a complex number denoted $\exp z$ and the rule $z \mapsto \exp z$ defines the **exponential function** of the complex variable z.

The exponential function transforms sums to products:

$$\exp(z + w) = (\exp z)(\exp w) \text{ for any complex } z \text{ and } w.$$

Indeed, due to the binomial formula, we have

$$(z + w)^n = \sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k} = n! \sum_{k+l=n} \frac{z^k}{k!} \frac{w^l}{l!}.$$

[3]Named after Étienne **Bézout** (1730–1783).

[4]Named after François Viéte (1540–1603) also known as Franciscus **Vieta**.

Rearranging the sum over all n as a double sum over k and l we get

$$\sum_{n=0}^{\infty} \frac{(z+w)^n}{n!} = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \frac{z^k}{k!}\frac{w^l}{l!} = (\sum_{k=0}^{\infty}\frac{z^k}{k!})(\sum_{l=0}^{\infty}\frac{w^l}{l!}).$$

The exponentials of complex conjugated numbers are conjugated:

$$\exp \bar{z} = \sum \frac{\bar{z}^n}{n!} = \overline{\sum \frac{z^n}{n!}} = \overline{\exp z}.$$

In particular, on the real axis the exponential function is real and coincides with the usual real exponential function $\exp x = e^x$ where $e = 1+1/2+1/6+...+1/n!+... = \exp(1)$. Extending this notation to complex numbers we can rewrite the above properties of $e^z = \exp z$ as $e^{z+w} = e^z e^w$, $e^{\bar{z}} = \overline{e^z}$.

On the imaginary axis, $w = e^{iy}$ satisfies $w\bar{w} = e^0 = 1$ and hence $|e^{iy}| = 1$. The way the imaginary axis is mapped by the exponential function to the unit circle is described by the following **Euler's formula**:

$$e^{i\theta} = \mathbf{cos}\,\theta + i\,\mathbf{sin}\,\theta.$$

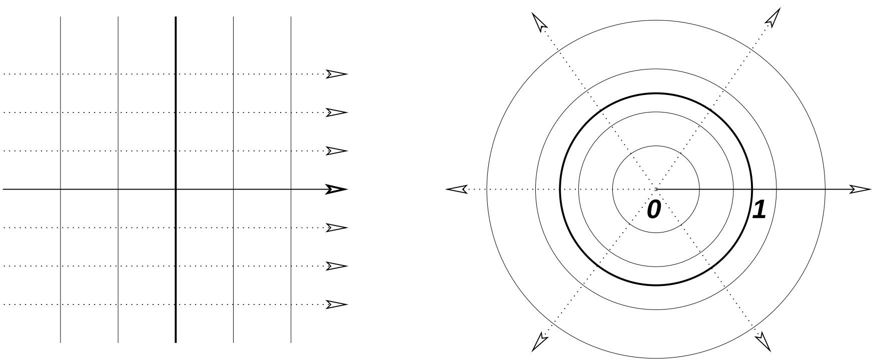

Figure 13 $z \longmapsto w = \exp z$

It is proved by comparison of $e^{i\theta} = \sum (i\theta)^n/n!$ with Taylor series for $\mathbf{cos}\,\theta$ and $\mathbf{sin}\,\theta$:

$$\begin{aligned} \mathrm{Re}\, e^{i\theta} &= 1 - \tfrac{\theta^2}{2} + \tfrac{\theta^4}{24} - ... = \sum (-1)^k \tfrac{\theta^{2k}}{(2k)!} = \mathbf{cos}\,\theta \\ \mathrm{Im}\, e^{i\theta} &= \theta - \tfrac{\theta^3}{6} + \tfrac{\theta^5}{120} - ... = \sum (-1)^k \tfrac{\theta^{2k+1}}{(2k+1)!} = \mathbf{sin}\,\theta \end{aligned}.$$

Thus $\theta \mapsto e^{i\theta}$ is the usual parameterization of the unit circle by the angular coordinate θ. In particular, $e^{2\pi i} = 1$ and therefore the exponential function is $2\pi i$-periodic: $e^{z+2\pi i} = e^z e^{2\pi i} = e^z$. Using Euler's formula we can rewrite the polar form of a non-zero complex number w as

$$w = |w| e^{i\,\mathbf{arg}\,w}.$$

EXERCISES

43. Prove the **Fundamental Formula of Mathematics** $e^{\pi i} + 1 = 0$ (which is so nicknamed, because it unifies the equality relation, and the operations of addition, multiplication, exponentiation with the famous numbers 0, 1, e, π, and i).

44. Represent $1 - i$ and $1 - i\sqrt{3}$ in the polar form $re^{i\theta}$.

45. Express $\mathbf{cos}\,\theta$ and $\mathbf{sin}\,\theta$ in terms of the complex exponential function. ϟ

46.* Show that the complex exponential function $w = \exp z$ maps the z-plane onto the entire w-plane except the origin $w = 0$, and describe geometrically the images on the w-plane of the lines of the coordinate grid of the z-plane.ϟ

47. Describe the image of the region $0 < \operatorname{Im} z < \pi$ under the exponential function $w = \exp z$. ✓

48. Compute the real and imaginary parts of the product $e^{i\phi}e^{i\psi}$, using Euler's formula, and deduce (once again) the addition formulas for $\mathbf{cos}(\phi + \psi)$ and $\mathbf{sin}(\phi + \psi)$.

49. Express the real and imaginary parts of $e^{3i\theta}$ in terms of $\mathbf{cos}\,\theta$ and $\mathbf{sin}\,\theta$, and deduce the triple argument formula for $\mathbf{cos}\,3\theta$ and $\mathbf{sin}\,3\theta$. ✓

50.* Prove the binomial formula $(z+w)^n = \sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k}$, where $\binom{n}{k} = n!/k!(n-k)!$. ϟ

3 A Model Example: Quadratic Curves

Conic Sections

On the coordinate plane, consider points (x, y), satisfying an equation of the form

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0.$$

Generally speaking, such points form a curve. The set of all solutions to the equation is called a **quadratic curve**, provided that not all of the coefficients a, b, c vanish.

Being a quadratic curve is a geometric property. Indeed, if the coordinate system is changed (say, rotated, stretched, or translated), the same curve will be described by a different equation, but the left-hand-side of the equation will remain a polynomial of degree 2.

Our goal in this section is to describe all possible quadratic curves geometrically (i.e. disregarding their positions with respect to coordinate systems); or, in other words, to *classify* quadratic equations in two variables up to suitable changes of the variables.

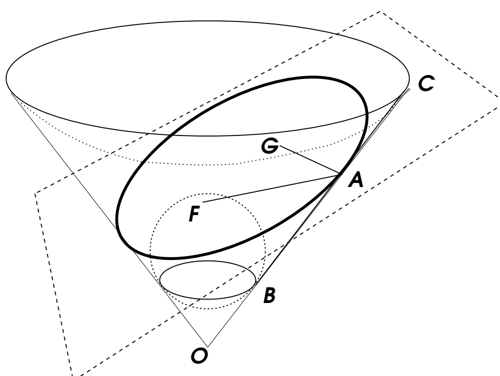

Figure 14

Example: **Dandelin's spheres.** The equation $x^2 + y^2 = z^2$ describes in a Cartesian coordinate system a cone (a half of which is shown on Figure 14). Intersecting the cone by planes, we obtain examples of quadratic curves. Indeed, substituting the equation $z = \alpha x + \beta y$ of a secting plane into the equation of the cone, we get a quadratic equation $x^2 + y^2 = (\alpha x + \beta y)^2$ (which actually describes the projection of the conic section to the horizontal plane).

The conic section on the picture is an **ellipse**. According to one of many equivalent definitions,[5] an ellipse consists of all points of the plane with a fixed sum of the distances to two given points (called **foci** of the ellipse). Our picture illustrates an elegant way[6] to locate the foci of a conic section.

Place into the conic cup two balls (a small and a large one), and inflate the former and deflate the latter until they touch the plane (one from inside, the other from outside). Then the points F and G of the tangency are the foci.

Indeed, let A be an arbitrary point on the conic section. The segments AF and AG lie in the cutting plane and are therefore tangent to the balls at the points F and G respectively. On the generatrix OA, mark the points B and C where it crosses the circles of tangency of the cone with the balls. Then AB and AC are tangent at these points to the respective balls. All tangent segments from a given point to a given ball have the same length. Hence we find that $|AF| = |AB|$, and $|AG| = |AC|$. Therefore $|AF| + |AG| = |BC|$. But $|BC|$ is the distance along the generatrix between two parallel horizontal circles on the cone, and is the same for all generatrices. We conclude that the sum $|AF| + |AG|$ stays fixed when the point A moves along our conic section.

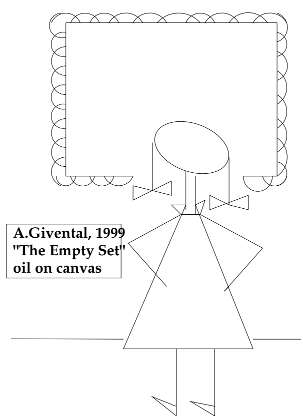

Figure 15. The Empty Set

Beside ellipses, we find among conic sections: **hyperbolas** (when a plane cuts through both halves of the cone), parabolas (cut by planes parallel to generatrices), and their degenerations (obtained

[5]According to a mock definition, "an ellipse is the circle inscribed into a square with unequal sides."

[6]Due to Germinal Pierre **Dandelin** (1794–1847).

when the cutting plane is replaced with the parallel one passing through the vertex O of the cone): just one point O, pairs of intersecting lines, and "double-lines." We will see that this list exhausts all possible quadratic curves, except two degenerate cases: pairs of parallel lines and (yes!) empty curves.

EXERCISES

51. Prove that a hyperbolic conic section consists of all points on the secting plane with a fixed *difference* of the distances to two points (called **foci**). Locate the foci by adjusting the construction of Dandelin's spheres.

52.* Prove that light rays emitted from one focus of an ellipse and reflected in it as in a mirror will focus at the other focus. Formulate and prove similar optical properties of hyperbolas and parabolas. ↯

53. Prove that the projections (Figure 16) of conic sections to the horizontal plane along the axis of the cone are quadratic curves.

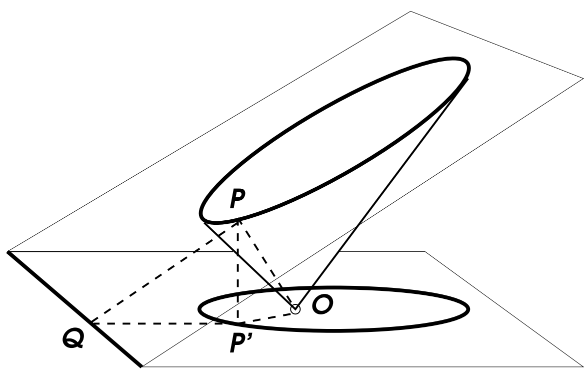

Figure 16

54.* Prove that the projections from the previous exercise can be characterized as plane curves formed by all points with a fixed ratio e (called **eccentricity**) between the distances to a fixed point (a **focus**) and a fixed line (called the **directrix**). ↯

55.* Show that $e > 1$ for hyperbolas, $e = 1$ for parabolas, and $1 > e > 0$ for ellipses (e.g. $e = |P'O|/|P'Q|$ in Figure 16). ↯

Orthogonal Diagonalization (Toy Version)

Let (x, y) be Cartesian coordinates on a Euclidean plane, and let Q be a **quadratic form** on the plane, i.e. a *homogeneous* degree-2 polynomial:

$$Q(x,y) = ax^2 + 2bxy + cy^2.$$

Theorem. ***Every quadratic form in a suitably rotated coordinate system assumes the form:***

$$Q = AX^2 + CY^2.$$

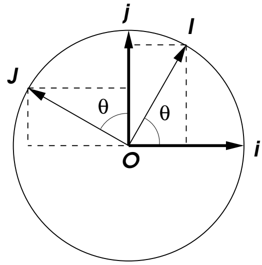

Figure 17

Proof. Rotating the unit coordinate vectors **i** and **j** counterclockwise through the angle θ (Figure 17), we obtain the following expressions for the unit coordinate vectors **I** and **J** of the rotated coordinate system:

$$\mathbf{I} = (\mathbf{cos}\,\theta)\mathbf{i} + (\mathbf{sin}\,\theta)\mathbf{j} \quad \text{and} \quad \mathbf{J} = -(\mathbf{sin}\,\theta)\mathbf{i} + (\mathbf{cos}\,\theta)\mathbf{j}.$$

Next, we express the radius-vector of any point in both coordinate systems:

$$x\mathbf{i} + y\mathbf{j} = X\mathbf{I} + Y\mathbf{J} = (X\,\mathbf{cos}\,\theta - Y\,\mathbf{sin}\,\theta)\mathbf{i} + (X\,\mathbf{sin}\,\theta + Y\,\mathbf{cos}\,\theta).$$

This shows that the old coordinates (x,y) are expressed in terms of the new coordinates (X,Y) by the formulas

$$x = X\,\mathbf{cos}\,\theta - Y\,\mathbf{sin}\,\theta, \quad y = X\,\mathbf{sin}\,\theta + Y\,\mathbf{cos}\,\theta. \qquad (*)$$

Substituting into $ax^2 + 2bxy + cy^2$, we rewrite the quadratic form in the new coordinates as $AX^2 + 2BXY + CY^2$, where A, B, C are certain expressions of a, b, c and θ. We want to show that choosing the rotation angle θ appropriately, we can make $2B = 0$. Indeed, making the substitution explicitly and ignoring X^2- and Y^2-terms, we find Q in the form

$$\cdots + XY\left(-2a\,\mathbf{sin}\,\theta\,\mathbf{cos}\,\theta + 2b(\mathbf{cos}^2\,\theta - \mathbf{sin}^2\,\theta) + 2c\,\mathbf{sin}\,\theta\,\mathbf{cos}\,\theta\right) + \dots$$

Thus $2B = (c-a)\,\mathbf{sin}\,2\theta + 2b\,\mathbf{cos}\,2\theta$. When $b = 0$, our task is trivial, as we can take $\theta = 0$. When $b \neq 0$, we can divide by $2b$ to obtain

$$\mathbf{cot}\,2\theta = \frac{a-c}{2b}.$$

Since **cot** assumes arbitrary real values, the theorem follows.

`Example.` For $Q = x^2 + xy + y^2$, we have $\mathbf{cot}\,2\theta = 0$, and find $2\theta = \pi/2 + \pi k$ $(k = 0, \pm 1, \pm 2, \dots)$, i.e. up to multiples of 2π, $\theta = \pm\pi/4$ or $\pm 3\pi/4$. (This is a general rule: together with a solution θ, the angle $\theta + \pi$ as well as $\theta \pm \pi/2$, also work. Could you give an *a priori* explanation?) Taking $\theta = \pi/4$, we compute $x = (X - Y)/\sqrt{2}$, $y = (X + Y)/\sqrt{2}$, and finally find:

$$x^2 + y^2 + xy = X^2 + Y^2 + \frac{1}{2}(X^2 - Y^2) = \frac{3}{2}X^2 + \frac{1}{2}Y^2.$$

EXERCISES

56. A line is called an **axis of symmetry** of a given function $Q(x, y)$ if the function takes on the same values at every pair of points symmetric about this line. Prove that every quadratic form has two perpendicular axes of symmetry. (They are called **principal axes**.) ϟ

57. Prove that if a line passing through the origin is an axis of symmetry of a quadratic form $Q = ax^2 + 2bxy + cy^2$, then the perpendicular line is also its axis of symmetry.ϟ

58. Can a quadratic form on the plane have > 2 axes of symmetry? ✓

59. Find axes of symmetry of the following quadratic forms Q:

$(a)\ x^2 + xy + y^2$, $(b)\ x^2 + 2xy + y^2$, $(c)\ x^2 + 4xy + y^2$.

Which of them have level curves $Q = const$ ellipses? hyperbolas? ✓

60. Transform the equation $23x^2 + 72xy + 2y^2 = 25$ to one of the standard forms by rotating the coordinate system explicitly. ϟ ✓

Completing the Squares

In our study of quadratic curves, the plan is to simplify the equation of the curve as much as possible by changing the coordinate system. In doing so we may assume that the coordinate system has already been rotated to make the coefficient at xy-term vanish. Therefore the equation at hands assumes the form

$$ax^2 + cy^2 + dx + ey + f = 0,$$

where a and c cannot both be zero. Our next step is based on **completing squares**: whenever one of these coefficients (say, a) is non-zero, we can remove the corresponding linear term (dx) this way:

$$ax^2 + dx = a(x^2 + \frac{d}{a}x) = a\left((x + \frac{d}{2a})^2 - \frac{d^2}{4a^2}\right) = aX^2 - \frac{d^2}{4a}.$$

Here $X = x + d/2a$, and this change represents translation of the origin of the coordinate system from the point $(x, y) = (0, 0)$ to $(x, y) = (-d/2a, 0)$.

`Example.` The equation $x^2 + y^2 = 2ry$ can be rewritten by completing the square in y as $x^2 + (y - r)^2 = r^2$. Therefore, it describes the circle of radius r centered at the point $(0, r)$ on the y-axis.

With the operations of completing the squares in one or both variables, renaming the variables if necessary, and dividing the whole equation by a non-zero number (which does not change the quadratic curve), we are well-armed to obtain the classification.

Classification of Quadratic Curves

Case I: $a \neq 0 \neq c$. The equation is reduced to $aX^2 + cY^2 = F$ by completing squares in each of the variables.

Sub-case (i): $F \neq 0$. Dividing the whole equation by F, we obtain the equation $(a/F)X^2 + (c/F)Y^2 = 1$. When both a/F and c/F are positive, the equation can be re-written as

$$\frac{X^2}{\alpha^2} + \frac{Y^2}{\beta^2} = 1.$$

This is the equation of an ellipse with **semiaxes** α and β (Figure 18). When one a/F and c/F have opposite signs, we get (possibly renaming the variables) the equation of a hyperbola (Figure 19)

$$\frac{X^2}{\alpha^2} - \frac{Y^2}{\beta^2} = 1.$$

When a/F and c/F are both negative, the equation has no real solutions, so that the quadratic curve is *empty* (Figure 15).

Sub-case (ii): $F = 0$. Then, when a and c have opposite signs (say, $a = \alpha^2 > 0$, and $c = -\gamma^2 < 0$), the equation $\alpha^2X^2 = \gamma^2Y^2$ describes a pair of intersecting lines $Y = \pm kX$, where $k = \alpha/\gamma$ (Figure 20). When a and c are of the same sign, the equation $aX^2 +$

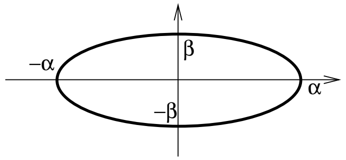

Figure 18. Ellipse

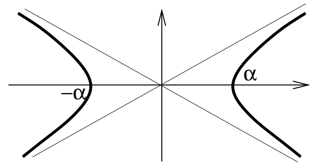

Figure 19. Hyperbola

$cY^2 = 0$ has only one real solution: $(X, Y) = (0, 0)$. The quadratic curve is a "thick" point.[7]

Case II: One of a, c ***is*** 0***.*** We may assume without loss of generality that $c = 0$. Since $a \neq 0$, we can still complete the square in x to obtain an equation of the form $aX^2 + ey + F = 0$.

Sub-case (i): $e \neq 0$***.*** Divide the whole equation by e and put $Y = y - F/e$ to arrive at the equation $Y = -aX^2/e$. This curve is a **parabola** $Y = kX^2$, where $k = -a/e \neq 0$ (Figure 21).

Sub-case (ii): $e = 0$***.*** The equation $X^2 = -F/a$ describes: a pair of parallel lines $X = \pm k$ (where $k = \sqrt{-F/a}$), or the empty set (when $F/a > 0$), or a "double-line" $X = 0$ (when $F = 0$).

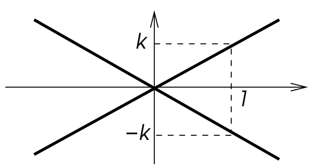

Figure 20

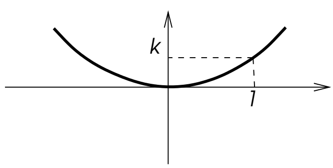

Figure 21

We have proved the following:

Theorem. ***Every quadratic curve on a Euclidean plane is one of the following: an ellipse, hyperbola, parabola, a pair of intersecting, parallel, or coinciding lines, a "thick" point or the empty set. In a suitable Cartesian coordinate system, the curve is described by one of the standard equations:***

$$\frac{X^2}{\alpha^2} \pm \frac{Y^2}{\beta^2} = 1, -1, \text{ \textbf{\textit{or}} } 0; \quad Y = kX^2; \quad X^2 = k.$$

[7]In fact this is the point of intersection of a pair of "imaginary" lines consisting of non-real solutions.

EXERCISES

61. Find the places of the following quadratic curves in our classification: $y = x^2 + x,\ xy = 1,\ xy = 0,\ xy = y,\ x^2 + x = y^2 - y,\ x^2 + x + y^2 - y = 0.$

62. Following the steps of our classification, reduce the quadratic equation $x^2 + xy + y^2 + \sqrt{2}(x - y) = 0$ to one of the standard forms. Show that the curve is an ellipse, and find its semiaxes. ϟ ✓

63. Use our classification theorem to prove that, with the exception of parabolas, each conic section has a center of symmetry. ϟ

64. Locate foci of (a) ellipses and (b) hyperbolas given by the standard equations $x^2/\alpha^2 \pm y^2/\beta^2 = 1$, where $\alpha > \beta > 0$. ✓

65. Show that "renaming coordinates" can be accomplished by a linear geometric transformation on the plane. ϟ

66. Prove that ellipses are obtained by stretching (or shrinking) unit circles in two perpendicular directions with two different coefficients.

67. From the Orthogonal Diagonalization Theorem on the plane, derive the following **Inertia Theorem** for quadratic forms in two variables: Every quadratic form on the plane in a suitable (but not necessarily Cartesian) coordinate system assumes one of the forms:

$$X^2 + Y^2,\ X^2 - Y^2,\ -X^2 - Y^2,\ X^2,\ -Y^2,\ 0.$$

Sketch graphs of these functions.

68. Complete squares to find out which of the following curves are ellipses and which are hyperbolas: ϟ ✓
$x^2 + 4xy = 1,\ x^2 + 2xy + 4y^2 = 1,\ x^2 + 4xy + 4y^2 = 1,\ x^2 + 6xy + 4y^2 = 1.$

69. Show that a quadratic form $ax^2 + 2bxy + cy^2$ is, up to a sign $\pm$, the square $(\alpha x + \beta y)^2$ of a linear function if and only if $ac = b^2$. ϟ

70. Show that if, in addition to rotation, reflection, and translation of coordinate systems, and multiplication of a quadratic equation by a non-zero constant, the change of scales of the coordinates is also allowed, then each quadratic equation can be transformed into one of the following 9 normal forms:

$$x^2 + y^2 = 1,\ x^2 + y^2 = 0,\ x^2 + y^2 = -1,\ x^2 - y^2 = 1,\ x^2 - y^2 = 0,$$
$$x^2 = y,\quad x^2 = 1,\quad x^2 = 0,\quad x^2 = -1.$$

71. Examine the curves defined by the above equations to conclude that they fall into 8 different types.

72. Find the place of the quadratic curve $x^2 - 4y^2 = 2x - 4y$ in the classification of quadratic curves. ✓

4 Problems of Linear Algebra

Classifications in Mathematics

Classifications are intended to bring order into seemingly complex or chaotic matters. Yet, there is a major difference between, say, our classification of quadratic curves and Carl **Linnaeus**' *Systema Naturae.*

For two quadratic curves to be in the same *class*, it is not enough that they share a number of features. What is required is a *transformation* of a prescribed type that would transform one of the curves into the other, and thus make them **equivalent** in this sense, i.e. *up to* such transformations.

What types of transformations are allowed (e.g., changes to *arbitrary* new coordinate systems, or only to *Cartesian* ones) may be a matter of choice. With every choice, the classification of objects of a certain kind (i.e. quadratic curves in our example) *up to* transformations of the selected type becomes a well-posed mathematical problem.

A complete answer to a classification problem should consist of
– a list of **normal** (or **canonical**) **forms**, i.e. representatives of the classes of equivalence, and
– a **classification theorem** establishing that each object of the kind (quadratic curve in our example) is equivalent to exactly one of the normal forms, i.e. in other words, that
(i) each object can be transformed into a normal form, and
(ii) no two normal forms can be transformed into each other.

Simply put, Linear Algebra deals with classifications of linear and/or quadratic equations, or systems of such equations. One might think that all that equations do is ask: *Solve us!* Unfortunately this attitude toward equations does not lead too far. It turns out that very few equations (and kinds of equations) can be explicitly *solved*, but all can be *studied* and many *classified.*

The idea is to replace a given "hard" (possibly unsolvable) equation with another one, the normal form, which should be chosen to be as "easy" as it is possible to find in the same equivalence class. Then the normal form should be studied (and hopefully "solved") thus providing information about the original "hard" equation.

What sort of information? Well, *any* sort that remains *invariant* under the equivalence transformations in question.

For example, in classification of quadratic curves up to changes of Cartesian coordinate systems, all equivalent ellipses are indistinguishable from each other *geometrically* (in particular, they have the same semiaxes) and differ only by the choice of a Cartesian coordinate system. However, if arbitrary rescaling of coordinates is also allowed, then all ellipses become indistinguishable from circles (but still different from hyperbolas, parabolas, etc.)

Whether a classification theorem really simplifies the matters, depends on the kind of objects in question, the chosen type of equivalence transformations, and the applications in mind. In practice, the problem often reduces to finding sufficiently simple normal forms and studying them in great detail.

The subject of linear algebra fits well into the general philosophy just outlined. Below, we formulate four model classification problems of linear algebra, solve them by bare hands in the simplest case of dimension 1, and state the respective general answers. Together with a number of variations and applications, which will be presented later in due course, these problems form what is usually considered the main course of linear algebra.

The Rank Theorem

`Question.` *Given m linear functions in n variables,*

$$\begin{array}{rcl} y_1 & = & a_{11}x_1 + ... + a_{1n}x_n \\ & ... & \\ y_m & = & a_{m1}x_1 + ... + a_{mn}x_n \end{array},$$

what is the simplest form to which they can be transformed by linear changes of the variables,

$$\begin{array}{rcl} y_1 & = & b_{11}Y_1 + ... + b_{1m}Y_m \\ & ... & \\ y_m & = & b_{m1}Y_1 + ... + b_{mm}Y_m \end{array}, \quad \begin{array}{rcl} x_1 & = & c_{11}X_1 + ... + c_{1n}X_n \\ & ... & \\ x_n & = & c_{n1}X_1 + ... + c_{nn}X_n \end{array} ?$$

`Example.` Consider a linear function in one variable: $y = ax$. We are allowed to make substitutions $y = bY$ and $x = cX$, where however $b \neq 0$ and $c \neq 0$ (so that we could reverse the substitutions). The substitutions will result in a new, transformed function: $Y = b^{-1}acX$. Clearly, if $a = 0$, then no matter what substitution we make, the linear function will remain identically zero. On the other hand, if $a \neq 0$, we can choose such values of b and c that the coefficient

$b^{-1}ac$ becomes equal to 1 (e.g. take $b = 1$ and $c = a^{-1}$). Thus, every linear function $y = ax$ is either identically zero: $Y = 0$, or can be transformed to $Y = X$.

Theorem. ***Every system of m linear functions in n variables can be transformed by suitable linear changes of dependent and independent variables to exactly one of the normal forms:***

$$Y_1 = X_1, \quad \ldots, \quad Y_r = X_r, \quad Y_{r+1} = 0, \quad \ldots, \quad Y_m = 0,$$

where $0 \le r \le m, n$***.***

The number r featuring in the answer is called the **rank** of the given system of m linear functions.

EXERCISES

73. Transform explicitly one linear function $y = -3x$ to the normal form prescribed by the Rank Theorem.

74. The same for the linear function $y = 3x_1 - 2x_2$.

75. The same for the system: $y_1 = x_1 + x_2$, $y_2 = x_1 - x_2$.

76. Prove that if two systems of m linear functions in n variables have the same rank, then they can be transformed into each other by linear changes of dependent and independent variables. ↯

The Inertia Theorem

Question. *Given a* **quadratic form** *(i.e. a homogeneous quadratic function) in n variables,*

$$Q = q_{11}x_1^2 + 2q_{12}x_1x_2 + 2q_{13}x_1x_3 + \ldots + q_{nn}x_n^2,$$

what is the simplest form to which it can be transformed by a linear change of the variables

$$\begin{array}{rcl} x_1 & = & c_{11}X_1 + \ldots + c_{1n}X_n \\ & \ldots & \\ x_n & = & c_{n1}X_1 + \ldots + c_{nn}X_n \end{array} \quad ?$$

Example. A quadratic form in one variable, x, has the form qx^2. A substitution $x = cX$ (with $c \neq 0$), transforms it into qc^2X^2. Of course, if $q = 0$, no substitution will change the fact that the function is identically zero. When $q \neq 0$, we can make the absolute value of coefficient qc^2 equal to 1 (by choosing $c = \pm\sqrt{|q^{-1}|}$). However, no substitution will change the sign of the coefficient (that is, a positive quadratic form will remain positive, and negative will remain negative). Thus, every quadratic form in one variable can be transformed to exactly one of these: X^2, $-X^2$, or 0.

`Theorem.` ***Every quadratic form in n variables can be transformed by a suitable linear change of the variables to exactly one of the normal forms:***

$$X_1^2 + ... + X_p^2 - X_{p+1}^2 - ... - X_{p+q}^2 \quad \textit{\textbf{where}}\ 0 \le p + q \le n.$$

Note that, in a way, the theorem claims that the n-dimensional case can be reduced to the sum (we will later call it "**direct sum**") of n one-dimensional answers found in the example: X^2, $-X^2$, or 0. The possibility of such reduction of a higher-dimensional problem to the direct sum of one-dimensional problems is a standard theme of linear algebra.

The numbers p and q of positive and negative squares in the normal form are called **inertia indices** of the quadratic form in question. If the quadratic form Q is known to be positive everywhere outside the origin, the Inertia Theorem tells us that in a suitable coordinate system Q assumes the form $X_1^2 + ... + X_n^2$, i.e. its inertia indices are $p = n,\ q = 0$.

EXERCISES

77. Transform explicitly the quadratic forms $4x^2$ and $-9y^2$ to their normal forms prescribed by the Inertia Theorem.

78. Transform the quadratic forms from the previous exercise into each other by a substitution $x = cy$ with possibly complex value of c.

79. Classify quadratic forms $Q = ax^2$ in one variable with *complex* coefficients (i.e. $a \in \mathbb{C}$) up to complex linear changes: $x = cX,\ c \in \mathbb{C}, c \neq 0$. ✓

80.* In the Inertia Theorem with $n = 2$, show that that there are six normal forms, and prove that they are pairwise non-equivalent. ϟ

81. Find the indices of inertia of the quadratic form $Q(x, y) = xy$. ϟ

82. Show that $X_1^2 + ... + X_n^2$ is the only one of the normal forms of the Inertia Theorem which is positive everywhere outside the origin.

83. Sketch the surfaces $Q(X_1, X_2, X_3) = 0$ for all normal forms in the Inertia Theorem with $n = 3$.

84. How many normal forms are there in the Inertia Theorem for quadratic forms in n variables? ✓

The Orthogonal Diagonalization Theorem

`Question.` *Given two homogeneous quadratic forms in n variables, $Q(x_1, ..., x_n)$ and $S(x_1, ..., x_n)$, of which the first one is known to be positive everywhere outside the origin, what is the simplest form to which they can be simultaneously transformed by a linear change of the variables?*

`Example.` In the case $n = 1$, we have $Q(x) = qx^2$, where $q > 0$, and $S(x) = sx^2$, where s is arbitrary. As we know, the first quadratic form is transformed by the substitution $x = q^{-1/2}X$ into X^2. The same transformation will change S into λX^2 with $\lambda = sq^{-1}$. Of course, one can make S to be $\pm X^2$ (if $s \neq 0$) by rescaling the variable once again, but this may destroy the form X^2 of the function Q. In fact the only substitutions $X = C\tilde{X}$ which *preserve* Q (i.e. don't change the coefficient) are those with $C = \pm 1$. Unfortunately such substitutions do not affect at all the coefficient λ in the function S: $\lambda X^2 = \lambda(\pm\tilde{X})^2 = \lambda\tilde{X}^2$. We conclude that each pair Q, S can be transformed into one of the pairs $X^2, \lambda X^2$, where λ is a real number, but two such pairs with different values of λ cannot be transformed into each other.

`Theorem.` ***Every pair Q, S of quadratic forms in n variables, of which Q is positive everywhere outside the origin, can be transformed by a linear change of the variables into exactly one of the normal forms***

$$Q = X_1^2 + ... + X_n^2, \ \ S = \lambda_1 X_1^2 + ... + \lambda_n X_n^2, \ \ \textit{\textbf{where}} \ \lambda_1 \geq ... \geq \lambda_n.$$

The real numbers $\lambda_1, \ldots, \lambda_n$ are called **eigenvalues** of the given pair of quadratic forms (and are often said to form their **spectrum**).

Note that this theorem, too, reduces the n-dimensional problem to the "direct sum" of n one-dimensional problems solved in our Example.

EXERCISES

85. Prove the Orthogonal Diagonalization Theorem for $n = 2$ using results of Section 3. ↯

86. Transform explicitly the quadratic form $Q = 3x^2 + 16y^2 + 9z^2$ to its normal form prescribed by the Inertia theorem, and apply the same transformation to the quadratic form $S = x^2 - 4y^2 + 12yz$.

87. Find the spectrum of the pair of quadratic forms: $Q = 3x^2 + 16y^2 + 9z^2$, $S = x^2 - 4y^2 + 12z^2$.

The Jordan Canonical Form Theorem

The fourth question deals with a system of n linear functions in n variables. Such an object is the special case of systems of m functions in n variables when $m = n$. According to the Rank Theorem, such a system of rank $r \leq n$ can be transformed to the form $Y_1 = X_1, \ldots, Y_r = X_r, Y_{r+1} = \cdots = Y_n = 0$ by linear changes of dependent and independent variables. There are many cases however where relevant information about the system is lost when dependent and independent variables are changed *separately*. This happens whenever both groups of variables describe objects in the same space (rather than in two different ones).

An important class of examples comes from the theory of Ordinary Differential Equations (ODE for short).

`Example.` Consider a linear first order ODE $\dot{x} = \lambda x$. It relates the values $x(t)$ of an unknown function, x, with its rate of change in time, $\dot{x}$ (which is the short notation for dx/dt). A rescaling of the function by $x = cX$ would make little sense if not accompanied with the simultaneous rescaling of the rate, $\dot{x} = c\dot{X}$ (we assume that the rescaling coefficient c is time-independent). Unfortunately, such a rescaling does not affect the form of the equation: $\dot{X} = c^{-1}\lambda c X = \lambda X$. We conclude that no two linear first order ODEs $\dot{x} = \lambda x$ with different values of the coefficient λ can be transformed into each other by a linear change of the variable.

We will describe the fourth classification problem in the context of the ODE theory, although it can be stated more abstractly as a problem about n linear functions in n variables, to be transformed by a single linear change acting on both dependent and independent variables *the same way.*

`Question.` *Given a system of n linear homogeneous 1st order constant coefficient ODEs in n unknowns:*

$$\begin{array}{rcl} \dot{x}_1 & = & a_{11}x_1 + \ldots + a_{1n}x_n \\ & \ldots & \\ \dot{x}_n & = & a_{n1}x_1 + \ldots + a_{nn}x_n \end{array},$$

what is the simplest form to which it can be transformed by a linear change of the unknowns:

$$\begin{array}{rcl} x_1 & = & c_{11}X_1 + \ldots + c_{1n}X_n \\ & \ldots & \\ x_n & = & c_{n1}X_1 + \ldots + c_{nn}X_n \end{array} \quad ?$$

There is an advantage in answering this question *over* $\mathbb{C}$, i.e. assuming that the coefficients c_{ij} in the change of variables, as well as the coefficients a_{ij} of the given ODE system are allowed to be complex numbers. The advantage is due to the unifying power of the Fundamental Theorem of Algebra, discussed in Supplement "Complex Numbers."

`Example.` Consider a single mth order linear ODE of the form:

$$(\frac{d}{dt} - \lambda)^m y = 0, \quad \text{where } \lambda \in \mathbb{C}.$$

By setting

$$y = x_1, \ \frac{d}{dt}y - \lambda y = x_2, \ (\frac{d}{dt} - \lambda)^2 y = x_3, \ \ldots, \ (\frac{d}{dt} - \lambda)^{m-1} y = x_m,$$

the equation can be written as the following system of m ODEs of the 1st order:

$$\begin{array}{lllllllll} \dot{x}_1 & = & \lambda x_1 & + & x_2 & & & & \\ \dot{x}_2 & = & & & \lambda x_2 & + & x_3 & & \\ & & & \cdots & & & & & \\ \dot{x}_{m-1} & = & & & & & \lambda x_{m-1} & + & x_m \\ \dot{x}_m & = & & & & & & & \lambda x_m \end{array} .$$

Let us call this system the **Jordan block** of size m with the eigenvalue λ. Introduce a **Jordan system** of several Jordan blocks of sizes $m_1, ..., m_r$ with the eigenvalues $\lambda_1, ..., \lambda_r$. It can be similarly compressed into the system

$$(\frac{d}{dt} - \lambda_1)^{m_1} y_1 = 0, \ ..., \ (\frac{d}{dt} - \lambda_r)^{m_r} y_r = 0$$

of r *unlinked* ODEs of the orders $m_1, ..., m_r$.

The numbers $\lambda_1, \ldots, \lambda_r$ here are not assumed to be necessarily distinct. In fact they are the roots of a certain degree n polynomial, $p(\lambda) = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_r)^{m_r}$, (called the **characteristic polynomial**), which can be associated with every linear ODE system, and does not change under the linear changes of the unknowns. When the polynomial has all its n roots distinct (that is, all $m_i = 1$, and $r = n$), the Jordan system assumes the form $\dot{x}_1 = \lambda_1 x_1, \ \ldots, \ \dot{x}_n = \lambda_n x_n$ of n unlinked first order ODEs discussed in our one-dimensional example. However, the theorem below implies that *not every* linear ODE system can be reduced to such a superposition (or direct sum) of one-dimensional ODEs. In particular, a single Jordan block of size $m > 1$ cannot be transformed into the superposition of one-dimensional ODEs.

`Theorem.` ***Every constant coefficient system of n linear 1st order ODEs in n unknowns can be transformed by a complex linear change of the unknowns into exactly one (up to reordering of the blocks) of the Jordan systems with $m_1 + ... + m_r = n$.***

EXERCISES

88. Find the general solution to the differential equation $\dot{x} = \lambda x$. ✓

89. Find the general solution to the system of ODE: $\dot{x} = 3x$, $\dot{y} = -y$, $\dot{z} = 0$. ✓

90. Verify that $y(t) = e^{\lambda t}\left(c_0 + tc_1 + ... + c_{m-1}t^{m-1}\right)$, where $c_i \in \mathbb{C}$ are arbitrary constants, is the general solution to the ODE $(\frac{d}{dt} - \lambda)^m y = 0$.

91. Rewrite the *pendulum* equation $\ddot{x} = -x$ as a system. ϟ

92.* Identify the Jordan form of the system $\dot{x}_1 = x_2$, $\dot{x}_2 = -x_1$. ✓

93.* Find the general solution to the system $\dot{x}_1 = x_2$, $\dot{x}_2 = 0$, and sketch the trajectories $(x_t(t), x_2(t))$ on the plane. Prove that the system cannot be transformed into any system $\dot{y}_1 = \lambda_1 y_1$, $\dot{y}_2 = \lambda_2 y_2$ of two unlinked ODEs.

Fools and Wizards

In the rest of the book we will undertake a more systematic study of the four basic problems and prove the classification theorems stated here. However, the reader (not unlike a fairy-tale hero) should be prepared to meet the following three challenges of the next Chapter.

Firstly, linear algebra has developed an adequate language, based on the abstract notion of **vector space**. It allows one to represent relevant mathematical objects and results in ways much less cumbersome and thus more efficient than those found in the previous discussion. This language is introduced at the beginning of Chapter 2. The challenge here is to get accustomed to the abstract way of thinking.

Secondly, one will find there much more diverse material than what has been described in the Introduction. This is because many mathematical objects and classification problems about them can be *reduced* (speaking roughly or literally) to the four problems discussed above. The challenge is to learn how to recognize situations where results of linear algebra can be helpful. Many of those objects will be introduced in the middle section of Chapter 2.

Finally, we will encounter one more fundamental result of linear algebra, which is not a classification, but an important (and beautiful) formula. It answers the question: *Which substitutions of the form*

$$\begin{aligned} x_1 &= c_{11}X_1 + ... + c_{1n}X_n \\ &\ ... \\ x_n &= c_{n1}X_1 + ... + c_{nn}X_n \end{aligned}$$

are indeed changes of the variables and can therefore be inverted by expressing $X_1, ..., X_n$ linearly in terms of $x_1, ..., x_n$, and how to describe such inversion explicitly? The answer is given in terms of the **determinant**, a remarkable function of n^2 variables $c_{11}, ..., c_{nn}$, which will also be studied in Chapter 2.

Let us describe now the principle by which our four main themes are grouped in Chapters 3 and 4.

Note that Jordan canonical forms and the normal forms in the Orthogonal Diagonalization Theorem do not form discrete lists, but instead depend on continuous parameters — the eigenvalues. Based on experience with many mathematical classifications, it is considered that the number of parameters on which equivalence classes in a given problem depend, is the right measure of complexity of the classification problem. Thus, Chapter 3 deals with **simple problems** of Linear Algebra, i.e. those classification problems where equivalence classes do not depend on continuous parameters. Respectively, the non-simple problems are studied in Chapter 4.

Finally, let us mention that the proverb: *Fools ask questions that wizards cannot answer,* fully applies in Linear Algebra. In addition to the four basic problems, there are many similarly looking questions that one can ask: for instance, to classify *triples* of quadratic forms in n variables up to linear changes of the variables. In fact, in this problem, the number of parameters, on which equivalence classes depend, grows with n at about the same rate as the number of parameters on which the three given quadratic forms depend. We will have a chance to touch upon such problems of Linear Algebra in the last, Epilogue section, in connection with *quivers.* The modern attitude toward such problems is that they are *unsolvable.*

EXERCISES

94. Using results of Section 3, derive the Inertia Theorem for $n = 2$.

95. Show that classification of real quadratic curves up to arbitrary linear inhomogeneous changes of coordinates consists of 8 equivalence classes. Show that if the coordinate systems are required to remain Cartesian, then there are infinitely many equivalence classes, which depend on 2 continuous parameters.

96. Is there any difference between classification of quadratic equations in two variables $F(x, y) = 0$ up to linear inhomogeneous changes of the variables and multiplications of the equations by non-zero constants, and the classification of quadratic curves, i.e. sets $\{(x, y)|F(x, y) = 0\}$ of solutions to such equations, up to the same type of coordinate transformations? ✓

97.[⋆] Derive the Orthogonal Diagonalization Theorem (as it is stated in this Section) in the case $n = 2$, using the "toy version" proved in Section 3, and *vice versa.* ϟ

98. Let us represent a quadratic form $ax^2 + 2bxy + cy^2$ by the point (a, b, c) in the 3-space. Show that the surface $ac = b^2$ is a cone. ϟ

99. Locate the 6 normal forms $(x^2 + y^2, x^2 - y^2, -x^2 - y^2, x^2, -y^2, 0)$ of the Inertia Theorem with respect to the cone $ac = b^2$ on Figure 9.

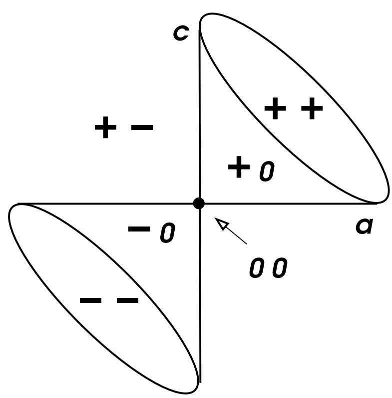

Figure 22

100. The cone $ac = b^2$ divides the 3-space into three regions (Figure 22). Show that these three regions, together with the two branches of the cone itself, and the origin form the partition of the space into 6 parts which exactly correspond to the 6 equivalence classes of the Inertia Theorem in dimension 2.

101. How many arbitrary coefficients are there in a quadratic form in n variables? ✓

102.[⋆] Show that equivalence classes of *triples* of quadratic forms in n variables must depend on at least $n^2/2$ parameters. ϟ